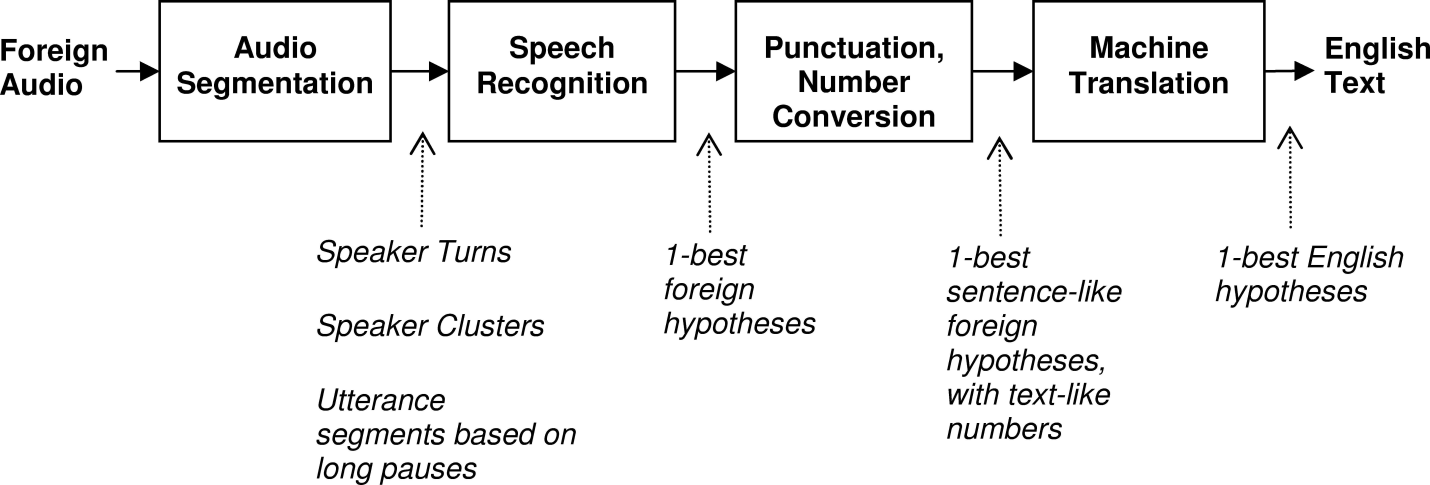
**Alaa Dukmak**

**120068042**

**Speech Recognition Systems (as part of MT)**

Speech recognition typically transfers spoken words to text. Speech recognition is a broad term which means it can recognize almost anybody's speech - such as a call-centre system designed to recognize many voices. As the majority of translation systems use phrase-based translation, it is important to maintain phrases and appropriate word-order. The choice of STT system combination approach for speech translation must take this into consideration.



**Voice Recognition Systems**

Voice recognition is the field of computer science that deals with designing computer systems that can recognize spoken words. Voice recognition systems are *speaker dependent.* Voice recognition implies only that the computer can take dictation, but it does not understandwhat is being said. Comprehending human languages falls under a different field of computer science called *natural language processing.* Because of their limitations and high cost, voice recognition systems have traditionally been used only in a few specialized situations.

**Areas of Applictions**

Speech [recognition](http://en.wikipedia.org/wiki/Recognition) applications include voice dialing, call routing, [domotic](http://en.wikipedia.org/wiki/Domotic" \o "Domotic) appliance control and content-based spoken audio search, simple data entry, preparation of structured documents, and speech-to-text processing, health care, Military, High-performance fighter aircraft, Helicopters, Battle management, Training air traffic controllers, Telephony and other domains, People with disabilities. Voice recognition systems are useful in instances when the user is unable to use a keyboard to enter data because his or her hands are occupied or disabled. Instead of typing commands, the user can simply speak into a headset. The software has been developed to provide a fast method of writing onto a computer and can help people with a variety of disabilities. It is useful for people with physical disabilities who often find typing difficult, painful or impossible. Voice recognition software can also help those with spelling difficulties, including users with dyslexic, because recognized words are always correctly spelled.

**Benefits**

Among the numerous benefits are: faster transportation to secretaries, enhanced clarity, enhanced typing accuracy, remote working flexibility, advanced workflow and ultimately, and faster turnaround of documentation.

**Limitations**

**Speech Recognition:**

The biggest limitation to speech recognition automating transcription is the software. The nature of narrative dictation is highly interpretive and often requires judgment that may be provided by a real human but not by an automated system. Another limitation has been the extensive amount of time required by the user or system provider to train the software.

**Voice Recognition:**

The voice recognition software will misunderstand some of the words spoken and it is necessary to correct the mistakes. When corrections are made the voice recognition software will adapt and learn. Accuracy should improve with careful dictation and correction.

**Provide examples of :**

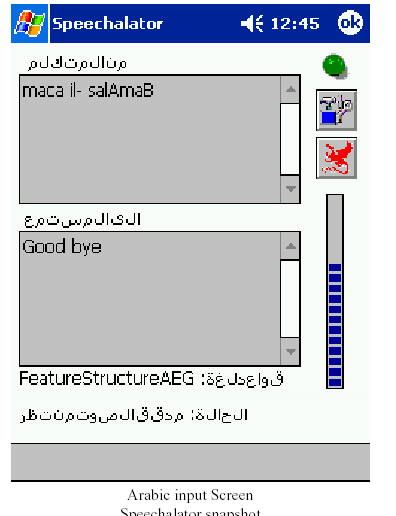
**One example of: text-to-speech online system(&features)**

|  |  |  |  |
| --- | --- | --- | --- |
| [**Active TTS Component 4.0.2010.401**](http://www.freedownloadscenter.com/Programming/ActiveX/Active_TTS_Component.html) http://www.freedownloadscenter.com/image/b_new.gif | | | |
| Shareware | $299.95 | 3-Apr-2010 | **1536K** |
| Directly convert text to wav, mp3, wma, ogg, vox, au and aiff format on the fly without temporary files created. Convert text to speech at high speed Support SAPI4 or SAPI5 compliant voices. | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **[Speak Aloud 2.0.2010.218](http://www.freedownloadscenter.com/Multimedia_and_Graphics/Misc__Sound_Tools/Speak_Aloud.html)** | http://www.freedownloadscenter.com/image/empty.gif | |  |  |  |  | | --- | --- | --- | --- | | [**Speak Aloud 2.0.2010.218**](http://www.freedownloadscenter.com/Multimedia_and_Graphics/Misc__Sound_Tools/Speak_Aloud.html) http://www.freedownloadscenter.com/image/b_new.gif | | | | | Shareware | $29.95 | 27-Feb-2010 | **4622K** | | Read text aloud with support for SAPI4 or SAPI5 compliant voices; Convert text to wav, mp3, wma, ogg, vox, au and aiff format on the fly, no temporary files needed, save your disk; Support batch conversion. | | | | |
|  |  |

**Arabic enabled speech recognition software:**

1. Sakhr’s ASR engine is the core technology recognizing spoken Arabic across a wide range of utterances, noisy environments, and Arabic accents from the Arabian Gulf, Egypt and Levant. This engine can be integrated with mobile client or telephony applications.
2. Part of the Babylon project has specific aspects to work with Arabic:



**Explore the area of: Speech-to-speech systems**

With increasing international travel and commerce, automatic speech-to-speech translation systems have attracted an interest from both theoretical and developmental communities. Automatic speech-to-speech translation is a highly complex task. A large amount of computation is involved to achieve reliable transformation performance. Resources are not just computation limited, but the memory and storage requirements, and the audio input and output requirements all tax current systems to their limits. For example, Sakhr S2S is a mobile device application that enables real-time verbal translation between English/Arabic, allowing an English or Arabic speaker to instantly understand and be understood by a speaker in the other language. Sakhr S2S combines Sakhr’s Arabic translation (MT), speech recognition (ASR), and text-to-speech technologies (TTS). Sakhr S2S is being provided to both commercial and government defense & intelligence markets.

**Major Speech Translations Systems**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | US | Europe | Japan | China |
| Research | Janus, DIPLOMAT, Tongues, Nespole! Maxtor | Verbmobil, Nespole! LC-Star, TC-Star | MATRIX | LodeStar,  Digital Olympics |
| Commercial | AT&T |  | NEC |  |
| Military | Phraselator, Babylon, LASER |  |  |  |

**One example of: speech-to-text online system(&features)**

Product Focus: Dragon NaturallySpeaking® v10

[[](http://www.pc-voice.co.uk/mainpages/products_software.asp#dns)](http://www.pc-voice.co.uk/mainpages/products_software.asp#dns)Dictate at 140 + words per minute in to your Windows application. Dragon NaturallySpeaking® voice recognition software enables users to work faster and more efficiently. Users can create documents, enter data, fill in forms, send e-mails, work on the Web, and control the desktop - all by voice.

**Some Example Programs**

There are several main voice recognition programs available:

**Dragon NaturallySpeaking**

This program is distributed by Nuance. NaturallySpeaking is recognized as the market leader.

The programs most commonly recommended by AbilityNet are:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Dragon NaturallySpeaking Preferred Version 10** | | | **Dragon Naturally Speaking Professional Version 10** | | |
| **Overview** | | A good choice for people with physical difficulties and dyslexia:  A comprehensive set of commands to allow voice control of the computer  Speech output facilities to help with corrections  It can be used with portable recorders | | A more advanced version for business users and for people who need to work completely hands-free:  Voice commands for Lotus Notes and Microsoft Outlook.  The ability to write new voice commands to assist easy hands-free use of the computer | |
| **Extra:**  **http://aramedia.com/asrb.jpgSakhr Software Launches Sakhr's Arabic ASR 4.0 Automatic Speech Recognition Engine**  Sakhr Software, the pioneering Arabic IT Company and supplier of speech recognition and text to speech technologies and business solutions, announced the availability of its engine Sakhr's ASR 4.0 software package. The version 4.0 of the ASR is the successor of Sakhr's ASR 3.0, containing new functionalities to ease application development and integration, increase accuracy and improve application usability. The ASR engine enables IVR systems to understand people's natural Arabic speech through spoken words in different Arabic accents. Developed for telephony speech applications, the recognition engine provides speech recognition and understanding for human-like telephony services. With ASR 4.0, developers can create robust and easy-to-use application interfaces designed to increase customer self-service, shorten transaction completion time and boost customer satisfaction levels. The ASR is designed to enable highly accurate, scalable, and easy to deploy voice solutions ranging from directory assistance, banking solutions, auto-attendant systems, product ordering, solutions for visually impaired and customer care to industry specific applications including airline reservations, brokerage, and many more. Sakhr's ASR software is a speaker-independent, open, standards-based recognition engine optimized for scalability, accuracy, and performance particularly for wireless environments and hands-free operation. The engine delivers unparalleled accuracy, often exceeding 95% in both wireless and landline deployments and exceptionally in noisy environments. It allows natural human-like dialogues through its robust grammar based barge-in feature. The latest ASR engine release builds on the advanced capabilities of the version 3.0 to provide memory and CPU savings and support for application vocabularies containing millions of entries."This version of the ASR is packaged with latest features required by the market and the highest accuracy ever achieved", said Jihad Salman, product manager for speech technologies at Sakhr. "Sakhr's ASR 4.0 allows the developers to easily speech-enable their telephony applications for the Arabic market and with its high accuracy enhances the quality of service while increasing the customers' satisfaction". | | | |

**References:**

<http://en.wikipedia.org/wiki/Speech_recognition>

<http://www.abilitynet.org.uk/content/factsheets/pdfs/Voice%20Recognition%20Software%20-%20An%20Introduction.pdf>

<http://www.bbn.com/resources/pdf/icassp07_STT_MT_Integration.pdf>

<http://www.freedownloadscenter.com/Search/text_to_speech.html>